

B. Loeb

#22

HMK

3/14/02



RAW SEQUENCE LISTING

DATE: 03/12/2002

PATENT APPLICATION: US/09/466,568

TIME: 14:46:00

Input Set : N:\Crf3\RULE60\09466568.raw

Output Set: N:\CRF3\03122002\I466568.raw

SEQUENCE LISTING

3 (1) GENERAL INFORMATION:

5 (i) APPLICANT: Crabtree, Gerald R.
 6 Schreiber, Stuart L.
 7 Spencer, David M.
 8 Wandless, Thomas J.
 9 Belshaw, Peter

11 (ii) TITLE OF INVENTION: REGULATED TRANSCRIPTION OF TARGETED
 12 GENES AND OTHER BIOLOGICAL EVENTS

14 (iii) NUMBER OF SEQUENCES: 81

16 (iv) CORRESPONDENCE ADDRESS:

17 (A) ADDRESSEE: ARIAD Pharmaceuticals, Inc.
 18 (B) STREET: 26 Landsdowne Street
 19 (C) CITY: Cambridge
 20 (D) STATE: Massachusetts
 21 (E) COUNTRY: USA
 22 (F) ZIP: 02139

24 (v) COMPUTER READABLE FORM:

25 (A) MEDIUM TYPE: Floppy disk
 26 (B) COMPUTER: IBM PC compatible
 27 (C) OPERATING SYSTEM: PC/DOS/MS/DOS
 28 (D) SOFTWARE: PatentIn Release #1.0, Version #1.25

30 (vi) CURRENT APPLICATION DATA:

C--> 31 (A) APPLICATION NUMBER: US/09/466,568
 C--> 32 (B) FILING DATE: 17-Dec-1999
 33 (C) CLASSIFICATION:

35 (vii) PRIOR APPLICATION DATA:

36 (A) APPLICATION NUMBER: US/09/157,753
 37 (B) FILING DATE:
 39 (A) APPLICATION NUMBER: 08/388,653
 40 (B) FILING DATE: 14-FEB-1995
 42 (A) APPLICATION NUMBER: US 08/478,386
 43 (B) FILING DATE: 07-JUN-1995

45 (viii) ATTORNEY/AGENT INFORMATION:

46 (A) NAME: Figg, E. Anthony
 47 (B) REGISTRATION NUMBER: 27,195
 48 (C) REFERENCE/DOCKET NUMBER: 2054-114A

50 (ix) TELECOMMUNICATION INFORMATION:

51 (A) TELEPHONE: (202) 783-6040
 52 (B) TELEFAX: (202) 783-6031

55 (2) INFORMATION FOR SEQ ID NO: 1:

57 (i) SEQUENCE CHARACTERISTICS:

58 (A) LENGTH: 14 amino acids

ENTERED

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/466,568

DATE: 03/12/2002

TIME: 14:46:00

Input Set : N:\Crif3\RULE60\09466568.raw

Output Set: N:\CRF3\03122002\I466568.raw

```

59          (B) TYPE: amino acid
60          (D) TOPOLOGY: linear
62      (ii) MOLECULE TYPE: peptide
65      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
67 Met Gly Ser Ser Lys Ser Lys Pro Lys Asp Pro Ser Gln Arg
68 1          5          10
71 (2) INFORMATION FOR SEQ ID NO: 2:
73      (i) SEQUENCE CHARACTERISTICS:
74          (A) LENGTH: 11 base pairs
75          (B) TYPE: nucleic acid
76          (C) STRANDEDNESS: single
77          (D) TOPOLOGY: linear
79      (ii) MOLECULE TYPE: cDNA
82      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
84 GTTAAGTTAA C 11
86 (2) INFORMATION FOR SEQ ID NO: 3:
88      (i) SEQUENCE CHARACTERISTICS:
89          (A) LENGTH: 11 base pairs
90          (B) TYPE: nucleic acid
91          (C) STRANDEDNESS: single
92          (D) TOPOLOGY: linear
94      (ii) MOLECULE TYPE: cDNA
97      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
99 TGACTCAGCG C 11
101 (2) INFORMATION FOR SEQ ID NO: 4:
103      (i) SEQUENCE CHARACTERISTICS:
104          (A) LENGTH: 33 base pairs
105          (B) TYPE: nucleic acid
106          (C) STRANDEDNESS: single
107          (D) TOPOLOGY: linear
109      (ii) MOLECULE TYPE: cDNA
111      (ix) FEATURE:
112          (A) NAME/KEY: misc_feature
113          (B) LOCATION: 6..11
114          (D) OTHER INFORMATION: /note= "Sac II restriction site."
116      (ix) FEATURE:
117          (A) NAME/KEY: misc_signal
118          (B) LOCATION: 12..16
119          (D) OTHER INFORMATION: /note= "Kozak sequence."
121      (ix) FEATURE:
122          (A) NAME/KEY: CDS
123          (B) LOCATION: 17..31
125      (ix) FEATURE:
126          (A) NAME/KEY: misc_feature
127          (B) LOCATION: 17..33
128          (D) OTHER INFORMATION: /note= "Region of homology with
129 target sequence."
132      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
134 CGACACCGCG GCCACC ATG GCC ACA ATT GGA GC 33

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/466,568

DATE: 03/12/2002

TIME: 14:46:00

Input Set : N:\Crf3\RULE60\09466568.raw

Output Set: N:\CRF3\03122002\I466568.raw

```

135             Met Ala Thr Ile Gly
136             1             5
139 (2) INFORMATION FOR SEQ ID NO: 5:
141   (i) SEQUENCE CHARACTERISTICS:
142       (A) LENGTH: 5 amino acids
143       (B) TYPE: amino acid
144       (D) TOPOLOGY: linear
146   (ii) MOLECULE TYPE: protein
148   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:
150 Met Ala Thr Ile Gly
151 1             5
154 (2) INFORMATION FOR SEQ ID NO: 6:
156   (i) SEQUENCE CHARACTERISTICS:
157       (A) LENGTH: 27 base pairs
158       (B) TYPE: nucleic acid
159       (C) STRANDEDNESS: single
160       (D) TOPOLOGY: linear
162   (ii) MOLECULE TYPE: cDNA
164   (ix) FEATURE:
165       (A) NAME/KEY: misc_feature
166       (B) LOCATION: 6..11
167       (D) OTHER INFORMATION: /note= "Xho I restriction site."
169   (ix) FEATURE:
170       (A) NAME/KEY: misc_feature
171       (B) LOCATION: 12..27
172       (D) OTHER INFORMATION: /note= "Region of homology with
173 target sequence."
176   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:
178 CGACACTCGA GAGCCCATGA CTTCTGG
181 (2) INFORMATION FOR SEQ ID NO: 7:
183   (i) SEQUENCE CHARACTERISTICS:
184       (A) LENGTH: 4 amino acids
185       (B) TYPE: amino acid
186       (C) STRANDEDNESS: single
187       (D) TOPOLOGY: linear
189   (ii) MOLECULE TYPE: peptide
192   (ix) FEATURE:
193       (A) NAME/KEY: Peptide
194       (B) LOCATION: 1..4
195       (D) OTHER INFORMATION: /note= "Translation product of
196 complement of SEQ ID NO:6, bases 9 to 20."
199   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:
201 Ser Trp Ala Leu
202 1
205 (2) INFORMATION FOR SEQ ID NO: 8:
207   (i) SEQUENCE CHARACTERISTICS:
208       (A) LENGTH: 41 base pairs
209       (B) TYPE: nucleic acid
210       (C) STRANDEDNESS: single

```

27

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/466,568

DATE: 03/12/2002

TIME: 14:46:00

Input Set : N:\Crif3\RULE60\09466568.raw

Output Set: N:\CRF3\03122002\I466568.raw

```

211         (D) TOPOLOGY: linear
213     (ii) MOLECULE TYPE: cDNA
215     (ix) FEATURE:
216         (A) NAME/KEY: misc_feature
217         (B) LOCATION: 6..11
218         (D) OTHER INFORMATION: /note= "Xho I restriction site."
220     (ix) FEATURE:
221         (A) NAME/KEY: misc_feature
222         (B) LOCATION: 12..41
223         (D) OTHER INFORMATION: /note= "Region of homology with
224 target sequence."
226     (ix) FEATURE:
227         (A) NAME/KEY: CDS
228         (B) LOCATION: 9..41
230     (ix) FEATURE:
231         (A) NAME/KEY: misc_feature
232         (B) LOCATION: 28
233         (D) OTHER INFORMATION: /note= "A to G."
236     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:
238 CGACACTC GAG CTC TGC TAC TTG CTA GGT GGA ATC CTC TTC      41
239     Glu Leu Cys Tyr Leu Leu Gly Gly Ile Leu Phe
240         1             5             10
243 (2) INFORMATION FOR SEQ ID NO: 9:
245     (i) SEQUENCE CHARACTERISTICS:
246         (A) LENGTH: 11 amino acids
247         (B) TYPE: amino acid
248         (D) TOPOLOGY: linear
250     (ii) MOLECULE TYPE: protein
252     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:
254 Glu Leu Cys Tyr Leu Leu Gly Gly Ile Leu Phe
255     1             5             10
258 (2) INFORMATION FOR SEQ ID NO: 10:
260     (i) SEQUENCE CHARACTERISTICS:
261         (A) LENGTH: 24 base pairs
262         (B) TYPE: nucleic acid
263         (C) STRANDEDNESS: single
W--> 287         (D) TOPOLOGY: /note= "Translational stop encoded
288             in complementary strand."
266     (ii) MOLECULE TYPE: cDNA
268     (ix) FEATURE:
269         (A) NAME/KEY: misc_feature
270         (B) LOCATION: 3..8
271         (D) OTHER INFORMATION: /note= "Eco RI restriction site."
273     (ix) FEATURE:
274         (A) NAME/KEY: misc_feature
275         (B) LOCATION: 9..24
276         (D) OTHER INFORMATION: /note= "Region of homology with
277 target sequence."
279     (ix) FEATURE:

```

RAW SEQUENCE LISTING

DATE: 03/12/2002

PATENT APPLICATION: US/09/466,568

TIME: 14:46:00

Input Set : N:\Crf3\RULE60\09466568.raw

Output Set: N:\CRF3\03122002\I466568.raw

```

280      (A) NAME/KEY: misc_feature
281      (B) LOCATION: 24
282      (D) OTHER INFORMATION: /note= "G to C."
284      (ix) FEATURE:
285          (A) NAME/KEY: misc_signal
286          (B) LOCATION: complement (9..11)
291      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:
293 GCGAATTCTT AGCGAGGGGC CAGC                                     24
295 (2) INFORMATION FOR SEQ ID NO: 11:
297     (i) SEQUENCE CHARACTERISTICS:
298         (A) LENGTH: 4 amino acids
299         (B) TYPE: amino acid
300         (C) STRANDEDNESS: single
301         (D) TOPOLOGY: linear
303     (ii) MOLECULE TYPE: peptide
306     (ix) FEATURE:
307         (A) NAME/KEY: Peptide
308         (B) LOCATION: 1..4
309         (D) OTHER INFORMATION: /note= "Translational product of
310 complement to SEQ ID NO:10, bases 12 to 23."
313     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:
315 Leu Ala Pro Arg
316 1
318 (2) INFORMATION FOR SEQ ID NO: 12:
320     (i) SEQUENCE CHARACTERISTICS:
321         (A) LENGTH: 33 base pairs
322         (B) TYPE: nucleic acid
323         (C) STRANDEDNESS: single
W--> 341         (D) TOPOLOGY: /note= "Translational stop signal
342             encoded on complementary strand."
326     (ii) MOLECULE TYPE: cDNA
328     (ix) FEATURE:
329         (A) NAME/KEY: misc_feature
330         (B) LOCATION: 3..8
331         (D) OTHER INFORMATION: /note= "Eco RI restriction."
333     (ix) FEATURE:
334         (A) NAME/KEY: misc_feature
335         (B) LOCATION: 12..17
336         (D) OTHER INFORMATION: /note= "Sal I restriction site."
338     (ix) FEATURE:
339         (A) NAME/KEY: misc_signal
340         (B) LOCATION: complement (9..11)
351     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:
353 GCGAATTCTT AGTCGACGCG AGGGGCCAGG GTC                                     33
355 (2) INFORMATION FOR SEQ ID NO: 13:
357     (i) SEQUENCE CHARACTERISTICS:
358         (A) LENGTH: 4 amino acids
359         (B) TYPE: amino acid
360         (C) STRANDEDNESS: single

```

RAW SEQUENCE LISTING ERROR SUMMARY

DATE: 03/12/2002

PATENT APPLICATION: US/09/466,568

TIME: 14:46:01

Input Set : N:\Crf3\RULE60\09466568.raw

Output Set: N:\CRF3\03122002\I466568.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:69; N Pos. 12,13,14,15,16,17,18,19,20

Seq#:70; Xaa Pos.3,4,5

Seq#:71; N Pos. 42,43,44,45,46,47,48,49,50

Seq#:72; Xaa Pos.13,14,15

Seq#:73; N Pos. 37,38,39,40,41,42,43,44,45

Seq#:74; Xaa Pos.9,10,11

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/466,568

DATE: 03/12/2002

TIME: 14:46:01

Input Set : N:\Crf3\RULE60\09466568.raw

Output Set: N:\CRF3\03122002\I466568.raw

L:31 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:32 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]
L:287 M:220 C: Keyword misspelled or invalid format, [(D) TOPOLOGY:]
L:287 M:246 W: Invalid value of Alpha Sequence Header Field, [TOPOLOGY:], SeqNo=10
L:341 M:220 C: Keyword misspelled or invalid format, [(D) TOPOLOGY:]
L:341 M:246 W: Invalid value of Alpha Sequence Header Field, [TOPOLOGY:], SeqNo=12
L:1435 M:220 C: Keyword misspelled or invalid format, [(2) INFORMATION FOR SEQ ID NO:]
L:1484 M:220 C: Keyword misspelled or invalid format, [(2) INFORMATION FOR SEQ ID NO:]
L:1670 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:70 after pos.:0
L:1702 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:72 after pos.:0
L:1734 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:74 after pos.:0